УДК 002.53:004.89

# ИСПОЛЬЗОВАНИЕ ОНТОЛОГИЙ В СИСТЕМАХ УПРАВЛЕНИЯ ЗНАНИЯМИ ОРГАНИЗАЦИЙ

С.В. Козлов, А.Ф. Тузовский, С.В. Чириков*, В.З. Ямпольский

Институт "Кибернетический центр" ТПУ, г. Томск
E-mail: TuzovskyAF@kms.cctpu.edu.ru
* ЗАО "ЭлеСи", г. Томск
E-mail: sergey.chirikov@elesy.ru

*Предложены варианты применения онтологий в качестве модели знаний организации для решения таких задач управления знаниями, как коммуникация, интеграция и выполнение логического вывода. Поясняется подход к построению и работе с онтологиями в системе управления знаниями.*

## Управление знаниями

Начиная с середины 90-х годов прошлого века, знания начали рассматриваться в промышленных компаниях в качестве важнейшего ресурса, ключевого фактора успеха и нового источника дохода [1]. Хотя компании управляли своими людскими и интеллектуальными активами задолго до этого, научная дисциплина "Управление знаниями" (УЗ) была сформирована именно в это время для решения таких специфических задач бизнеса. К сожалению, не существует общепринятого определения понятия "Управление знаниями", что в основном связано с различными точками зрения специалистов на понимание термина "знание". В данной статье под УЗ понимается комплексный набор мероприятий, направленных на поддержание в организации системного порядка работы с информационно-знаниевыми ресурсами и специалистами для поиска, накопления и облегчения доступа к знаниям, повторного или многократного их использования [2]. Как показывает опыт ведущих компаний, это позволяет организациям обучаться более эффективно и максимизировать получаемый эффект от индивидуальных и коллективных источников информации и знаний. Общепринято, что УЗ оперирует сложной смесью проблем, касающихся людей, организаций и технологий. Понимание этой сложности является важным для разработки технических решений специалистами, которые, по крайней мере, осведомлены о других аспектах дисциплины.

В момент появления в производственной сфере дисциплины УЗ, множество исследований в области Искусственного Интеллекта (ИИ) были направлены на решение проблем представления знаний и выполнения логического вывода на знаниях, которые ***объединяли ранний опыт концептуального моделирования с использованием формальных логик***. Центральным элементом этих исследований являлось понятие "онтология", которое представляет собой совместно используемое, формальное описание некоторой предметной области. И хотя, исследователи онтологий рассматривают УЗ, как одну из наиболее важных областей их применения [3], онтологии все еще слабо используются в разработке и использовании систем УЗ в бизнесе.

В первые годы становления УЗ, информационные технологии рекламировались как средство решения всех проблем управления знаниями. Разработчики спешно старались переименовать программные системы по управлению информацией (базы данных по проектам, базы данных по лучшим практическим решениям, инсталляции Lotus Notes и т.п.) в решения по УЗ. Вторая волна проникновения ИТ на этот рынок включала успешные решения в проектах по интеграции крупных баз данных и по бизнес-интеллекту (*business intelligence*). Третий, текущий этап, уделяет основное внимание предоставлению пользователям, сотрудникам и потребителям удобного доступа к информации, посредством интранет, порталов и т.п. Появилось программное обеспечение, предоставляющее технологии, которые поддерживают процессы и человеческие аспекты УЗ: управление бизнес-процессами, рабочие пространства для совместной работы, инструменты управления проектами и т.п.

Сегодня, можно найти сотни продуктов и сервисов призванных поддерживать ту или иную функцию УЗ. Несмотря на разработку все более сложных методов поддержки управления активами знаний и процессами работы с ними, следует отметить, что в производственной среде имеется неудовлетворенность результатами использования информационных технологии в прошлом. Однако, из анализа этих неудач, возникло понимание того, что УЗ по своей природе является чем-то большим, чем установка какого-либо нового программного обеспечения.

Из анализа литературы и программных приложений по УЗ можно сделать следующий вывод: хотя и можно найти прототипы приложений по УЗ, основанные на онтологиях (в литературе по онтологиям), очень мало литературных источников по УЗ даже упоминает использование онтологий. Те авторы, которые это делают, обычно описывают их в разделах "будущие технологии УЗ". То есть, большинство источников предполагает, что онтологии еще не вышли за пределы научных исследований, и поэтому в настоящее время еще нет возможности их реального использования.

## Варианты использования онтологий

Для пояснения пользы от использования онтологий в УЗ, требует поиск таких процессов и сценариев работы со знаниями, которые могут напрямую дать выигрыш от применения онтологий. Для

этого рассмотрим более тщательно основные свойства онтологий. *Онтология, исходя из наиболее часто используемого определения в области ИИ, является – совместно используемой, формальной классификацией предметной области* [4]. Онтологии являются моделями данных, обладающих двумя специфическими особенностями, которые ведут к понятию совместного понимания или семантике:

1. *Онтологии строятся на основе совместного понимания предметной области в рамках сообщества.* Это понимание представляется соглашением экспертов по поводу понятий и отношений, которые имеются в предметной области.
2. *Онтологии используют способ представления, который может обрабатываться компьютерными программами* (т.е., записываются с использованием формальных языков, таких как RDFS или OWL [5]), что дает возможность компьютерам работать с онтологиями. К таким действиям относятся передача онтологий между компьютерами, хранение онтологий, проверка согласованности онтологий, выполнение логических выводов на онтологиях и с помощью онтологий.

Эти особенности дают возможность поднять уровень интеллектуальности программных систем, путем включения в них знаний о предметной области в форме онтологий. В то время как компьютерные программы могут манипулировать и выполнять логический вывод на знаниях о предметной области с помощью онтологии, понимание того, что обозначают эти символы и правила, остается функцией человека. Поэтому онтология не может существовать без сообщества, которое ее поддерживает. Онтология сплетает понимание предметной области человеком и компьютером. Предполагается, что совместно используемое понимание представленное онтологиями может быть использовано для управления тремя ключевыми процессами УЗ [6]: (1) ***Коммуникация***; (2) ***Интеграция*** и (3) ***Логический вывод***. Эти процессы основываются друг на друге и отражают увеличивающиеся уровни формальности в смысле использованных онтологий, по мере возрастания сложности требуемых знаний. В связи с этим, увеличиваются требования к уровню формальности онтологий, начиная с простых словарей до формальных логических структур, которые позволяют выполнять сложные логические выводы. Далее рассмотрим более детально выделенные варианты применения онтологий в системах УЗ.

***(1) Использование онтологий для коммуникации***. Онтологии облегчают коммуникацию и общение, предоставляя совместно используемые понятия, которые могут использоваться для формулировки и записи утверждений (запросов, ответов и т.п.) об экземплярах предметной области. Для получателя таких сообщений, онтология помогает понять сообщение и избежать двусмысленности, предоставляя контекст для их интерпретации. В общем случае, использование общепринятого, хорошо понятного языка может помочь организациям лучше обучаться, становиться более осведомленными, и быстрее реагировать на изменение окружающей среды. При согласованном применении, совместно используемый и формально признанный язык способствует тому, что рабочие группы (например, разные группы работы с клиентами) могут работать более согласованно, увеличивается понимание друг друга. Это использование онтологий может быть расширено по масштабу до коммуникации между отделами и обмена сообщениями с партнерами. Единственной проблемой в этом случае является то, чтобы тщательно разработанная и строгой онтологии не ограничивать творчество сотрудников и их желание экспериментировать.

***(2) Использование онтологий для Интеграции разнородных и распределенных ресурсов информации и знания.*** Онтологии являются более чем простыми словарями точно определенных понятий: их реальная сила заключается в описании связей между сущностями предметной области. Можно сказать [7], что в некотором отношении, весь мир может рассматриваться, как множество различных взаимосвязей. Любой словарь, считается хранилищем смысла, но он определяет слова только по их отношению к другим словам. Элемент информации в действительности определяется тем, с какими другими элементами и как он связан.

Полноценные онтологии имеют больше возможностей для описания отношений, чем простые таксономические иерархии, которые имеются в современных решениях по УЗ. Отношения (связи) между понятиями значительно увеличивают количество способов осуществлять навигацию и поиск по предметной области, выполнять анализ, классификацию и визуализацию знаний.

Для того, чтобы использовать эту богатую паутину понятий для интеграции информации, разнородные информационные и знаниевые ресурсы могут быть размечены с помощью терминов онтологии. Этот процесс, также называемый *семантической нормализацией*, разделяет на группы концептуально связанные элементы информации, независимо от их форматов и представлений. Хотя эти действия подобны присоединению ключевых слов или классификационных кодов к информации, результирующие описания имеют точную интерпретацию, предоставляемую онтологией.

Термин "ресурс" здесь используется в самом широком смысле, включая базы данных, текстовый контент (документы), профили компетентности пользователей, web-сервисы и т.п. В приложениях УЗ на предприятиях наиболее многообещающей разработкой является область "Корпоративной памяти", так как этим системам обычно требуется работать с разнородными и неструктурированными ресурсами знаний, включающие документы, e-mail сообщения, web-страницы, календари, заметки с совещаний и т.п. Разнородность затрудняет обнаружение взаимосвязей ("знание о том, что мы знаем"), а неструктурированные форматы

(спроектированные для чтения людьми) не позволяют выполнять поиск программно. Фактически, эта технология является особенно привлекательной для описания мультимедиа, такого как изображения и видео фрагменты, содержание которых напрямую недоступно компьютеру.

Положительным здесь является то, что система получает выигрыш от более точного, основанного на понятиях поиска и просмотра. Онтологии предметных областей, используемые для классификации информации, могут быть использованы для создания интеллектуальных интерфейсов, которые помогут пользователям в формулировке запросов в терминах онтологии.

Сценарий интеграции размещен над уровнем коммуникации. В этом случае, коммуникация происходит между создателем онтологии и описаний (метаданных) и действующим лицом, которое выполняет поиск с помощью онтологии. В большинстве случаев, эта коммуникация опосредуется компьютерной системой, которая помогает пользователю формировать их запросы, предлагая альтернативные термины, обобщения, уточнения, и т.п., основываясь на понимании компьютером языка онтологии. Улучшенный доступ к коллективным знаниям организации ведет к более высокой производительности работников, так как они (повторно) используют существующие знания более эффективно и защищают себя от информационной перегрузки.

***(3) Использование онтологий для выполнения логического вывода.*** Выполнение логического вывода над терминологией является наиболее сложным вариантом использования онтологий. В то время как два раннее рассмотренных уровня использования онтологий построены на знаниях о том, какие есть виды объектов в исследуемой предметной области (коммуникация), и как они связаны (интеграция), данный вариант использования требует знаний о том, почему они связаны. Другими словами, этот сценарий касается правил и принципов, которые стоят за некоторой концептуализацией.

Для комплексных типов приложений, таких как классификация (которые напрямую поддерживаются Дескриптивной Логикой), новый стандарт языка онтологий OWL DL позволяет автоматически находить правильное место для понятия (например, такого как "описание продукции") в иерархии описаний. Этот сервис может использоваться в ходе разработки онтологий, например, для автоматического расположения вновь созданных комплексных понятий, а также при вводе данных или при ответе на сложные запросы, сформулированные как частично определенные понятия.

### Построение онтологии для систем УЗ

В связи с важностью использования онтологий в системах УЗ возникает задача создания и сопровождения онтологий, которая является нетривиальной, и которой отводится важная роль в комплексе мероприятий по созданию систем УЗ. На сегодняшний день разработан ряд методологических подходов к проектированию онтологий: скелетная модель Uschold и King [8], методология Gr?ninger и Fox [9], METHONTOLOGY [10], методология проекта On-To-Knowledge (OTK) [11] и т.д.

Многие методологии ориентированы на создание онтологий для решения конкретных задач (проекты Enterprise Ontology и Toronto Virtual Enterprise, проект On-To-Knowledge). Одной из наиболее целостных и обоснованных методологий на сегодняшний день является методология OTK (рис. 1), разработанная в рамках Европейского исследовательского проекта (IST-1999-10132).

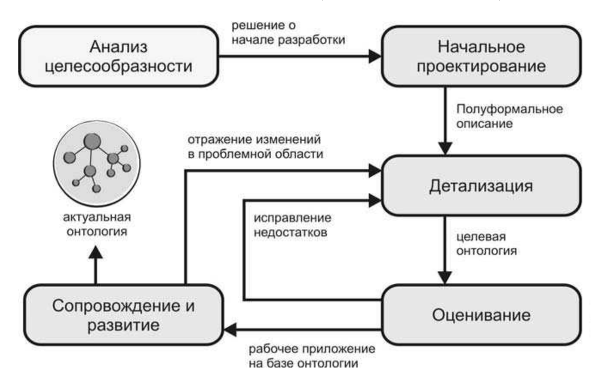



***Рис. 1.*** *Этапы создания онтологии*

На основе анализа этих методологий предложен подход к разработке онтологий состоящий из следующих этапов:

**1. Анализ целесообразности.** Этот этап имеет подготовительный характер. Необходимо выявить проблемы работы со знаниями в организации, для решения которых может быть использована программная система на основе онтологии. Выделяются приоритетные направления, которые должны обеспечиваться подобной системой. Решение о целесообразности создания и внедрения системы является необходимым условием для выполнения следующих этапов.

**2. Начальное проектирование.** На этапе начального проектирования составляется спецификация требований и создается прототип онтологии. Спецификация требований содержит: цель создания онтологии; описание предметной области и оценку масштаба онтологии; описание программных средств, использующих онтологию; список источников знаний используемых для построения онтологии; список потенциальных пользователей и описание сценариев работы с онтологией. Результатом этапа начального проектирования является неформализованное описание онтологии.

**3. Детализация.** На этапе детализации неформализованное описание онтологии служит основой для создания модели. Этот этап включает сбор данных и извлечение знаний из источников определенных в спецификации. В качестве источников могут быть использованы документы, книги, диаграммы, таблицы, должностные инструкции, эксперты в предметной области, уже структурирован-

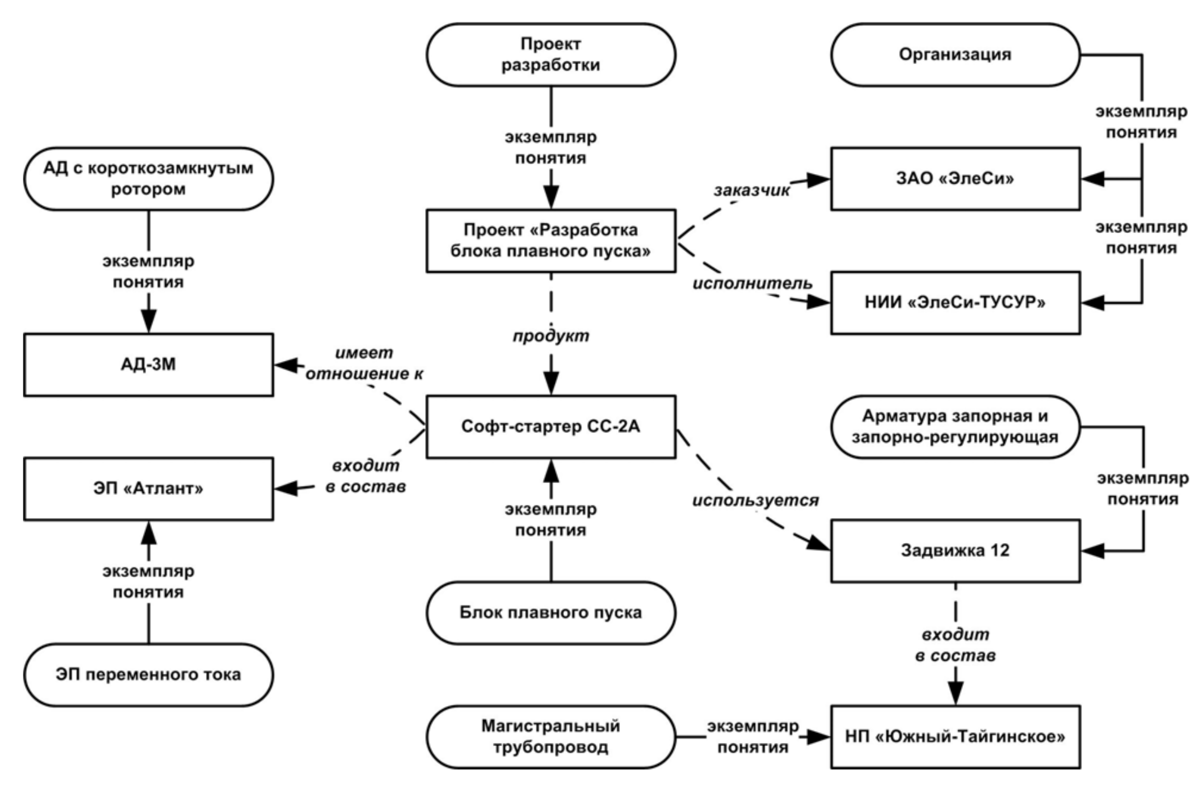



***Рис. 2.*** *Отношения между экземплярами понятий онтологии*

ная информация - существующие общие онтологии, словари, тезаурусы и др. Для извлечения знаний применяются различные методы: мозговой штурм, интервью, анализ текста, специализированные программные средства. Пример онтологического описания применительно к сфере деятельности компании "ЭлеСи" представлен на рис. 2.

Также на этапе детализации выполняется описание онтологии при помощи формальных языков представления знаний, например, таких как RDFS или OWL. Выбор языка зависит от задач, которые планируется решать при помощи онтологии. Это связано с тем, что формальные языки представления знаний отличаются своими выразительными возможностями и уровнем поддержки логического вывода.

В методологиях проектирования онтологий обычно выделяют два этапа проектирования: создание неформализованного описания онтологии; создания формальной спецификации на языке представления знаний. Наличие этих двух описаний – важная особенность онтологий. Неформализованное описание добавляется к формальной модели в виде комментариев.

Для формализации онтологии используются специализированные редакторы. Примерами таких редакторов являются Protégé и OilEd. Редактор онтологий Protégé 3.0 позволяет инженеру по знаниям визуально конструировать онтологию, не задумываясь о том, как она будет храниться в памяти компьютера.

**4. Оценивание.** На этапе оценивания необходимо определить, удовлетворяет ли разработанная онтология спецификации и могут ли использующие онтологию приложения обеспечить заявленный уровень компетенции. Для этого необходимо протестировать прототип онтологии, обладающей базовой функциональностью. Предыдущий этап детализации является итеративным и тесно связан с этапом оценивания онтологии. Если анализ онтологии выявляет недостатки реализации, этап детализации повторяется вновь.

**5. Сопровождение и развитие.** Разработанная онтология не должна рассматриваться как набор статичных, неизменных во времени понятий и терминов. С течением времени значение терминов предметной области может изменяться, термины могут получать новый смысл в связи с появлением новых и развитием существующих технологий. Чтобы эффективно выполнять свои функции, онтология предметной области должна поддерживаться в актуальном состоянии.

После создания приложения на базе онтологии, проект переходит на этап сопровождения и развития. Это, в основном, организационный процесс, для которого четко регламентированы все процедуры изменения онтологии, а также отвечающие за их выполнение лица. Развитие онтологии рекомендуется осуществлять в виде последовательных этапов накопления изменений, тестирования и анализа возможных эффектов от модификации с последующим переходом к новой версии.

Описанные варианты использования онтологий опробованы в рамках выполнения договора по теме "Разработка проекта и базовых элементов си-

стемы управления знаниями компании ЭлеСи" сотрудниками института "Кибернетический Центр" ТПУ совместно со специалистами компании "ЭлеСи". Для этой цели было разработано онтологическое описание предметной области "Управляемый электропривод", являющееся частью онтологии "Автоматизация". Принципы создания данной онтологии были аналогичны методологии ОТК. Онтология описывает понятия и отношения области знаний по управляемым электроприводам (578 понятий и 15 отношений).

Разработанная онтология применялась для систематизации знаний, используемых в компании "ЭлеСи": в качестве терминологической базы для описания источников знаний; в качестве основы для семантического поиска источников знаний (сотрудников и документов); в качестве элемента создаваемого семантического web портала компании.

### Заключение

На основе проведенного анализа можно сделать вывод о том, что в настоящее время имеются три эффективных способа использоваться онтологий при разработке и функционировании систем управления знаниями: коммуникация, интеграция и логический вывод. В результате практического опыта было выявлено, что наиболее эффективным является использование онтологий для интеграции распределенных и разнородных источников информации и знаний. Использование онтологий позволяет организовать навигацию и семантический поиск информации и знаний, что не возможно реализовать в ранее используемых подходах к проектированию и работе систем управления знаниями, не основанных на единой формальной модели знаний организации.

### СПИСОК ЛИТЕРАТУРЫ


1. Davenport T., Prusak L. Working Knowledge: how organizations manage what they know. – Boston: Harvard Business School Press, 1998. – 200 p.
2. Тузовский А.Ф., Чириков С.В., Ямпольский В.З. Системы управления знаниями (методы и технологии). – Томск: Изд-во НТЛ, 2005. – 260 с.
3. Leger A. Successful Scenarios for Ontology-based Applications V1.0. OntoWeb: Deliverable 2.2. – Paris, 2002. – 100 p. [Электронный ресурс]. Режим доступа: http://www.aifb.uni-karlsruhe.de/WBS/ysu/publications/OntoWeb_Del_2-1.pdf
4. Gruber T. A translation approach to portable ontology specifications // Knowledge Acquisition. – 1993. – V. 5. – № 2. – P. 199–220.
5. Abecker A., Ludger van Elst. Handbook on Ontologies. – Berlin: Springer Verlag, 2003. – P. 436–454.
6. Mika P., Akkermans H. Towards a New Synthesis of Ontology Technology and Knowledge Management: Technical Report IR-BI-001. – Amsterdam, 2004. – 37 p. [Электронный ресурс]. Режим доступа: http://www.cs.vu.nl/~pmika/research/papers/IR-BI-001.pdf
7. Berners-Lee T., Fischetti M., Dertouzos M. Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web by its Inventor. – San Francisco: Harper San Francisco, 1999. – 240 p.
8. Uschold M., King M. Towards a Methodology for Building Ontologies // Workshop on Basic Ontological Issues in Knowledge Sharing. – Montreal, 1995. [Электронный ресурс]. Режим доступа: http://www.aiai.ed.ac.uk/project/pub/documents/1995/95-ont-ijcai95-ont-method.ps
9. Grüninger M., Fox M.S. Methodology for the Design and Evaluation of Ontologies // Proc. of IJCAI'95. – Montreal, 1995. [Электронный ресурс]. Режим доступа: http://citeseer.ist.psu.edu/grninger95methodology.html
10. Fernández-López M., Gomez-Perez A., Juristo N. METHONTOLOGY: From Ontological Art Towards Ontological Engineering // Workshop on Ontological Engineering. – Stanford: Spring Symposium Series, 1997. – P. 33–40.
11. Staab S., Schnurr H.P., Studer R., Sure Y. Knowledge processes and ontologies. // IEEE Intelligent Systems – 2001. – V. 16. – № 1. – P. 26–34.
12. The Protégé Ontology Editor and Knowledge Acquisition System. [Электронный ресурс]. Режим доступа: http://protege.stanford.edu/